

WHAT IS CLAIMED IS:

1. A split model network driver comprising:
a host module driver provided on a host, the host connected to a switched fabric;
an I/O unit module driver on a remote I/O unit, the I/O unit connected to the switched fabric;
wherein both the host module driver and the I/O unit module driver include a push-push messaging layer to communicate messages across the switched fabric using message Sends and RDMA writes.
2. The split model network driver of claim 1 wherein the host and the I/O unit each includes a virtual interface (VI) work queue pair.
3. The split model network driver of claim 2 wherein the VI work queue pair comprises a send queue and a receive queue.
4. The split model network driver of claim 1 wherein both the host module driver and the I/O unit module driver include a push-push messaging layer to communicate messages across the switched fabric using message Sends and RDMA writes, the RDMA write including data and a transfer indication to identify a buffer consumed at a destination.

5. The split model network driver of claim 4 wherein the transfer indication is provided as immediate data in the RDMA write.

6. The split model network driver of claim 1 wherein the messaging layers of the host module driver and the I/O unit module driver communicate control and configuration messages over a first channel and communicate data over a second channel.

7. The split model network driver of claim 1 wherein the messaging layers of the host module driver and the I/O unit module driver communicate buffer management messages over a first channel and communicate data over a second channel.

8. An apparatus comprising:

- a first virtual interface (VI) work queue pair;
- a second virtual interface (VI) work queue pair;
- a channel adapter coupled to the VI work queue pairs to interface the apparatus to a switched fabric;
- a messaging layer to communicate buffer management messages over the switched fabric using the first work queue pair, and to communicate RDMA write messages over the switched fabric using the second VI work queue pair, one or more of the RDMA write messages including a transfer indication within the RDMA write message that identifies a buffer consumed by the RDMA write.

9. The apparatus of claim 8 wherein the apparatus comprises a host unit.

10. The apparatus of claim 8 wherein the apparatus comprises an I/O unit.

11. The apparatus of claim 8 wherein the messaging layer is a push-push-messaging layer.

12. The apparatus of claim 8 wherein the messaging layer is provided to communicate buffer management messages over a first channel using the first work queue pair, and to communicate RDMA write messages over a second channel using the second VI work queue pair.

13. The apparatus of claim 8 wherein the transfer indication is provided as immediate data in the RDMA write.

14. An apparatus comprising:

a virtual interface (VI) work queue;

a host channel adapter coupled to the VI work queue, the channel adapter to interface the host to a switched fabric;

a push-push messaging layer to communicate control and buffer management messages over a first channel and to communicate RDMA write messages over a second channel, one or more of the RDMA write messages

including a transfer indication within the RDMA write message, the transfer indication identifying a buffer that is consumed at a destination by a RDMA write.

15. The apparatus of claim 14 wherein the apparatus comprises a host and the channel adapter comprises a host channel adapter.

16. The apparatus of claim 14 wherein the apparatus comprises an I/O unit and the channel adapter comprises a target channel adapter.

17. The apparatus of claim 16 wherein the apparatus further comprises a network interface controller (NIC) to interface the I/O unit to another network, the control channel being used by the messaging layer of the I/O unit to receive NIC configuration messages received across the switched fabric.

18. A host comprising:
at least one virtual interface (VI) work queue pair;
a host channel adapter coupled to the VI work queue pair, the channel adapter to interface the host to a switched fabric;
a push-push messaging layer to communicate control and buffer management messages over a first channel and to communicate RDMA write messages over a second channel, the messaging layers communicating a transfer indication message over the first channel after an RDMA write message

and separate from the RDMA write message, the transfer indication message identifying a buffer that was consumed at a destination by a RDMA write.

19. A method comprising:

establishing a connected channel between first and second nodes across a switched fabric by associating a work queue pair of the first node with a work queue pair of the second node;

communicating data from the first node to the second node using a push-push RDMA write to a pre-registered buffer at the second node, the RDMA write also including a transfer indication identifying the buffer at the second node that received the data.

20. The method of claim 19 wherein the transfer indication is provided as immediate data within the RDMA write.

21. The method of claim 19 and further comprising establishing a second connected channel between the first and second nodes, the second channel to communicate control and buffer management messages between the nodes.

22. A computer program, embodied in tangible medium, when executed causes a computing device to:

communicate data from the first node to the second node using a push-push RDMA write to a pre-registered buffer at the second node, the RDMA write also including a transfer indication identifying the buffer at the second node that received the data.

THE UNIVERSITY OF CHICAGO